**Objectives Documentation**

**1.** Data cleaning

* First of by looking at the Hospital ER data, we can clearly see patient\_sat\_score has *null* values. I observed the patient\_sat\_score have 0 values as well. So, I replaced the null values with 0 value. So that it makes it easier for us to get the proper insights from our data.
* Secondly, I observed the columns Appointment fees and Total bill has Format Currency which is correct but while looking at these columns it looks like any other column of numbers. To visualize better, I entered dollar sign to make it look like a currency.

**2.** Assess the Average Waiting Time:

* To calculate the average waiting time for patients, we can create a measure for the average waiting time. Here’s a DAX measure which I used to get the waiting time: *Average waiting time = AVERAGE('Hospital ER'[patient\_waittime (in Minutes)])*
* As the problem demands a numerical value to be calculated, I have shown a **Card** mentioning Average waiting time in title.

A screen shot of a number

Description automatically generated

**3.** Visits by Department Referral:

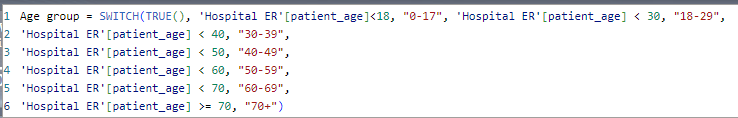
* To calculate which departments are most frequently visited based on the total number of visits to each department, I have used a stacked bar chart with the 'Department Referral' on the x-axis and the count of visits on the y-axis.
* It is clearly visible from the below chart that the department ‘General Practice’ has the highest number of visits and ‘Renal’ with the lowest number of visits.

A screenshot of a computer

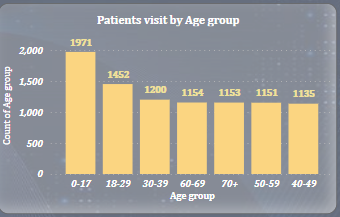
Description automatically generated

**4.** Patient Visits by Age Group:

* To segregate patients’ visits according to different age groups, I have applied DAX formula SWITCH on patient\_age column. For instance, if patient\_age<18, the age group will be “0-17.”
* DAX:



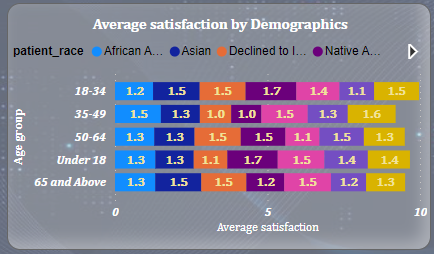
* To visualize, I have dragged a bar chart where x-axis represents Age group and y-axis represents number of patients visited, Please refer the below attached snap of the chart.



**5.** Average Satisfaction by Demographics:

*Analysis:*

* We can calculate the relationship between patient satisfaction scores, their age group, and racial backgrounds by creating a new calculated column to group patients into different age groups. I have used DAX formula *SWITCH* to generate the age group column.
* On the visualization part, I have two choices between Scatter plot or stacked bar chart. The choice depends on the specific aspects of the data we want to emphasize and the insights we want to derive. Therefore, I have used stacked bar chart to represent the average satisfaction score of patients along with the patient’s age group and the Patient’s race.
* Scatter plot is used to explore the relationship between two continuous variables (e.g., satisfaction scores and age) and additionally we can use colour to represent a categorical variable (e.g., patient race).
  + Advantages: Provides a visual representation of the distribution, concentration, and potential trends in satisfaction scores across different age groups and racial backgrounds.
* Bar chart is used when we want to compare the average satisfaction scores across different categories (e.g., age groups and racial backgrounds).
  + Advantages: Allows for easy comparison between different categories. It's effective for highlighting differences in average satisfaction scores and identifying patterns or trends.
* To visualize better stacked bar chart looks easier to interpret the results with the help of data labels. Please find below the attached chart.

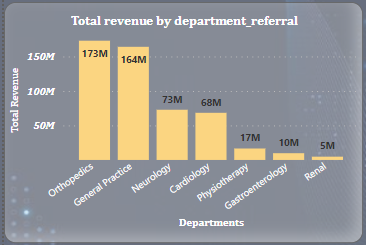


*Interpretation:*

* The average satisfaction score including all the races and age group is 1.36.
* Most consistent average\_sat\_score is observed in Patient\_race: Asian and White.
* The lowest average\_sat\_score is 1.0 and same for the highest is 1.7.
* Age group 35-49 has the lowest average\_sat\_score which means this age group is the area of improvement for patient’s experience. If we observe two races has the average of 1.0 which is the lowest among all the scores.

**6.** Revenue of each department:

* We can evaluate revenue of each department by using Group By feature to group the data by ‘department\_referral’.
* Summarize the total revenue for each department by aggregating the ‘Total Bill’ column.
* I have used stacked column chart to show the contribution of different revenue within each department. Please refer the below snap for the same.



**7.** Relation between number of visits by Patient Gender:

* I’ve used Pie chart to represent the number of visits based on patient gender.

A yellow circle with black numbers and red dots

Description automatically generated

* The distribution which is showed in the above chart in the pie chart indicates the percentage of hospital visits by patient gender.
* Male (51.05%):
  + Approximately 51.05% of the hospital visits in your dataset are attributed to male patients.
  + This suggests that males make up a slightly higher proportion of the patient population in the hospital during the analysed period.
* Female (48.69%):
  + About 48.69% of the hospital visits are attributed to female patients.
  + Females represent a slightly lower but significant portion of the patient population compared to males.
* NC (0.26%):
  + The category 'NC' (which likely stands for 'Not Categorized' or another designation) comprises a very small percentage of the total visits (0.26%).
  + It's crucial to investigate why there are patients not categorized by gender. This could be due to data entry issues or specific cases where gender information is not available.
* Interpretation:
  + The distribution between male and female patients is relatively balanced, with a slight majority of visits from males.
  + The presence of a small percentage in the 'NC' category indicates the need for further investigation into the completeness and accuracy of gender data in your dataset.

**8.** Which department is charging the highest appointment fees in general?

* We simply require the highest appointment fees overall, not department-specific, therefore I made a new measure and utilized the DAX function MAX () to get the highest fees from the appointment fees column.
* I have used a card visual to show the highest appointment fees in the report as you can see in the below snap.
* DAX:

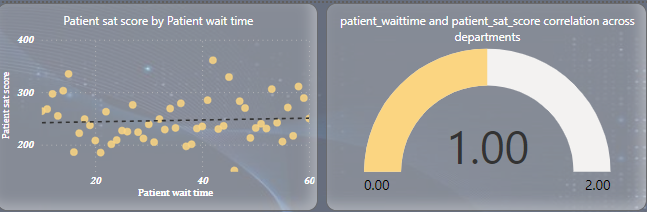
Highest appointement fees = MAX(hospital\_doctor\_data[Appointment Fees])



**Subjective Documentation**

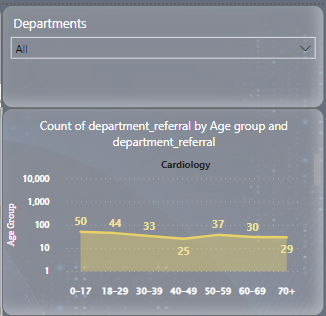
**Q.** What is the relation between patient wait time and satisfaction scores?

* I have used scatter plot to show the relationship between patient wait time and satisfaction scores.
* To enhance the plot, I have added a trendline to visually represent the overall trend in the data.
* I have also added a Guage chart which displays correlation coefficient between wait time and sat score. The correlation coefficient is 1, which indicates a perfect positive linear relationship between these two variables. It means that there is a perfect positive correlation between patient wait time and satisfaction scores.
* This suggests that as patient wait time increases, satisfaction scores also increase, and vice versa.
* Here’s the snap of the plot and gauge chart.



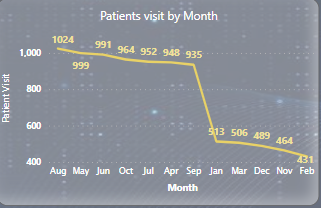
**Q.** How do patient demographics affect the frequency of visits to different departments?

* Depending on the data, I have used Line chart visualization for analysing the frequency of visits to different departments based on patient demographics.
* For interaction, I have used **Field parameter** when we select a specific department and see how it affects the other elements of the chart.
* Conclusion:
  + One noticeable pattern in the data which is observed infie every department we can see is the age group: 0-17 is the highest number of patients visiting frequently.
  + Secondly, the patient visit in “General practice” department is the highest among all other departments and the lowest patient visit has been observed in the “Renal” department.
  + The age group 40-49 is the dropping point where patient visits are going down in almost every department. Please refer below snap of the visualization.



**Q.** Is there a noticeable trend in the volume of patient visits throughout the year?

* I have used Line charts to show the volume of patients visiting throughout the year. As you can see in the below chart, I have added data labels to the chart for better clarity.
* Across all 12 months, the number of patients visiting has ranged from 431 to 1,024.
* There is a seasonal trend which is being followed here -
  + The number of patients was low during winters. The lowest number of patients observed was in the month of February.
  + Since the spring season, the number of patients drastically increased, At 1,024 August month had the highest number of patients visiting frequency and was 137.59% higher than February which was the lowest.
  + ﻿August month accounted for 11.11% of number of patient.



**Q.** Which age group reports the highest and lowest satisfaction scores?

* We require two measures to find out the highest and lowest satisfaction scores based on the age group, the first one will be age group and second one will be Average satisfaction scores. I have already calculated these measures previously.
* I have created a bar chart to visualize the results as we can see in the snapshot below. Also, I have added data label to identify the satisfaction scores for each age groups.
* At **1.41**, Age group 18-29 records the highest average satisfaction scores.
* At **1.28**, Age group 70+ records the lowest average satisfaction scores.

A graph of numbers and a number of people

Description automatically generated with medium confidence

**Q.** The hospital management intends to offer discounts to patients. How should these bonuses be assigned to patients, on what basis?

* We must analyze patient satisfaction scores; total bill amounts and demographic information to identify criteria for assigning discounts.
* First, we will check if the patients are eligible for discounts. We can create a measure which will evaluate patient sat score should be less than 5. If true, then patient is eligible for discounts. If false, then patient is not eligible for discounts.
* I have created a column and used DAX formula SWITCH.

DAX:

Discount Level = SWITCH(

TRUE(),

'Hospital ER'[Patient Sat Score] < 3, "High Discount",

'Hospital ER'[Patient Sat Score] < 6, "Medium Discount",

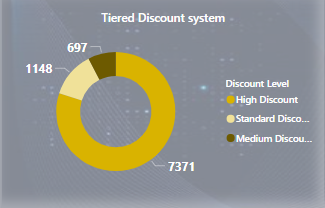
"Standard Discount")

* Here’s the visuals where I have used Pie chart to show eligibility of the number of patients for discounts.

A yellow circle with a black line

Description automatically generated

* Secondly, we will implement a tiered discount system based on the specific criteria to ensure fairness and targeted assistance. It involves different discount levels for different satisfaction scores.
* To visualize the number of patients by discount level, I’ve used donut chart, with the Data labels outside the chart for better clarity. Please check the snapshot of the chart below.



**Q.** The hospital has a budget to hire 2-3 new doctors, In which departments they should hire?

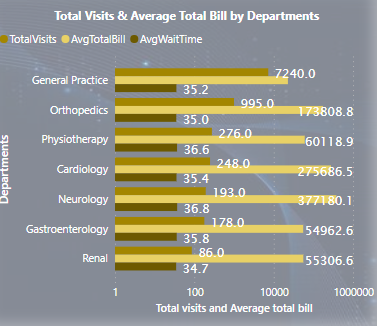
* This task involves providing suggestions on which departments the hospital should consider for hiring 2-3 new doctors.
* Firstly, we should create key metrics that will help in deciding where to hire new doctors. I have created few measures which will help us to understand the scenario of hiring new doctors.
* First measure will evaluate the count of all the patients visiting, second measure will evaluate the Average of total bill of Patient’s every single visit and the third measure will identify the departments with longer wait times, as these may need additional staffing so, it will evaluate the average waiting time for the patient.
* Here’s the DAX measures for all three metrics:

Total Patient Visits: TotalVisits = COUNTROWS('hospital\_doctor\_data')

Average Total Bill per Visit: AvgTotalBill = AVERAGE('hospital\_doctor\_data'[Total Bill])

Average Wait Time: AvgWaitTime = AVERAGE('hospital\_doctor\_data'[Patient Wait Time])

* To visualize patient load and revenue by department, I have shown a clustered bar chart as you can see in the below snapshot. I have added all three metrics with each department.

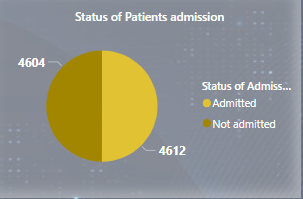


**Q**. Is the hospital profitable?

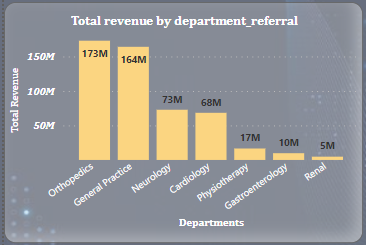
* To see whether the hospital is profitable or not, I will first evaluate the overall revenue.
* I will create a DAX measure to estimate the overall revenue of the hospital.

DAX: OverallRevenue = SUM('hospital\_doctor\_data'[Total Bill])

* As the data related to expense is very low, we can see in the below snap, patient admin flag as True which contributes 50.04% that is 4612 patients were admitted to the hospital to all the departments.



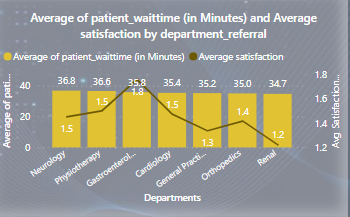
* If we assume the expense based on the services provided to these patients and we are rolling out discounts as well to the patients.
* Analysze revenue by department:
  + We can evaluate revenue of each department by using Group By feature to group the data by ‘department\_referral’.
  + Summarize the total revenue for each department by aggregating the ‘Total Bill’ column.
  + I have used stacked column chart to show the contribution of different revenue within each department. Please refer the below snap for the same.



* Conclusion: Based on the above analysis, the hospital’s overall financial health seems stable. We have not explored any additional sources of revenue, such as special services, partnerships, or other income-generating activities. Henceforth, we can say the hospital is in profitable state.

**Q.** Any department for which the waiting time is oddly large?

* To identify any department for which the waiting time is oddly large, I have analyzed the waiting time for each department.
* I have used a Line and clustered column chart to display the waiting time for each department and used a line to highlight the average satisfaction score.
* As we can see in the below chart, we can see the patient visit in Gastroenterology is the third highest department with the number of patients visiting but the waiting time for the patients is unusually high. Surprisingly, the average satisfaction score for this department is the highest among all the satisfaction scores.
* Reason for high waiting time:
  + There could be potential reasons behind this waiting time, it could be related to resource constraints or operational issues, as the data is not available so we cannot drill down to the concerned areas to investigate.
* Recommendations:
  + There can improve the process which is being followed and try to propose a new process to follow.
  + If it is required resource allocation can be changed as well to have a smooth process to follow for patients.



**Q.** Come up with the strategies to provide discounts to the patients.

* By visualizing and analysing patient demographics in relation to satisfaction scores and total bills, we can gain valuable insights that will inform targeted discount strategies, ensuring fairness and relevance for different patient groups.
* As we have already developed Tiered discount system previously, now I have explored the strategy in Patient’s satisfaction scores with patient ethnic backgrounds.
* I have used Line chart to visualize the strategies which can be implemented for patients. I have identified a trend that may influence discount strategy.
* Conclusion:
  + If we observe the below line chart, we can see the race “Native American/Alaska Native” generates the highest Average total bill with the second highest Average satisfaction score. That is, they are providing a good satisfaction score and generating a good revenue for Hospital. Hence, hospital should consider rolling out discounts for these patients.
* Recommendations:
  + We can develop special discount programs for specific demographic groups.
  + Regularly review the effectiveness of discount strategies based on demographics.
  + Be open to adjustments and refinements based on ongoing analysis and feedback.

